

AD_____

Award Number: W81XWH-06-1-0057

TITLE: Identification and Characterization of an X-Linked Familial Prostate Cancer Gene

PRINCIPAL INVESTIGATOR: Brian Yaspan, Ph.D.

CONTRACTING ORGANIZATION: Vanderbilt University Medical Center
Nashville TN 37203-6917

REPORT DATE: November 2006

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</small>					
1. REPORT DATE (DD-MM-YYYY) 01/11/06		2. REPORT TYPE Annual Summary		3. DATES COVERED (From - To) 24 Oct 05 – 23 Oct 06	
4. TITLE AND SUBTITLE Identification and Characterization of an X-Linked Familial Prostate Cancer Gene				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-06-1-0057	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Brian Yaspan, Ph.D. E-Mail: brian.l.yaspan@vanderbilt.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Vanderbilt University Medical Center Nashville TN 37203-6917				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT None provided.					
15. SUBJECT TERMS None provided.					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			USAMRMC
			UU	8	19b. TELEPHONE NUMBER (include area code)

Table of Contents

Introduction	4
Body.....	4
Key Research Accomplishments.....	7
Reportable Outcomes.....	8
Conclusions.....	8
References.....	none

Introduction

There is a significant heritable component of prostate cancer. Increased familial relative risk is observed across multiple populations (European, Asian-American, African-American and Caucasian). Male first degree relatives of prostate cancer patients have a two- to three-fold increased risk. Segregation analyses support genetic rather than shared environmental risk. Twin cancer concordance studies reveal a higher heritable risk for prostate cancer than for any other common cancer. A Scandinavian study of 44,788 twin pairs estimated that 42% risk of prostate cancer is heritable; the concordance rate was 21% among monozygotic twins, and 6% among dizygotic twins. Large-scale US twin studies have revealed comparable concordance rates. Additional epidemiological studies have been consistent with X-linked transmission, identifying higher risk for a man with an affected brother relative to one with an affected father.

Despite the overwhelming genetic predisposition evidence, the identification of prostate cancer susceptibility genes has been difficult. Linkage studies have resulted in the identification of several loci difficult to confirm across study populations. However, summary studies of genome-wide scans for prostate cancer susceptibility loci in general confirm two loci, HPC-1 and HPC-X.

Our study seeks to identify a candidate gene or genes conferring prostate cancer susceptibility at locus HPC-X in a US Caucasian study population. We hypothesize that a gene or genes at HPC-X harbor common moderate-penetrance variants predisposing to prostate cancer, with a role much greater than that inferred through study of rare HPC families. We looked at shared haplotypes in founder populations and found two intervals likely to harbor prostate cancer susceptibility genes. We have chosen to first focus on one interval at locus HPC-X (termed HPC-X region A) due to shared haplotype association evidence in the founder populations of Finland, Iceland and Ashkenazim.

Body

Accomplishments

Task 1. To analyze the genetic architecture of locus HPC-X and to determine the haplotype(s) associated with disease, both protective and risk. (Months 1-12):

- a. Perform *de novo* SNP discovery at predicted or known genes and derive a set of survey SNPs spanning the HPC-X locus and a density of 3-5 kb from dbSNP
- b. Genotype population for polymorphism
- c. Analyze genotypes to determine genetic architecture of HPC-X

Task 1a-1c was completed this year. To derive a set of survey SNPs spanning the interval at HPC-X we assayed SNPs found in dbSNP for polymorphism in a subset of 41 probands from our US Caucasian population. We identified 203 polymorphic SNPs at a density of 2.4 kb/SNP. *De novo* SNP discovery methods were performed via SSCP gel

techniques in the same set of 41 probands. Gels were run at room temperature, and if a polymorphism was not found, the gel was repeated at 4° C. *De novo* SNP discovery added 66 additional SNPs for a total of 269 SNPs. This set of SNPs was assayed in half of our population of 486 prostate cancer probands and 453 screened cancer-free controls for haplotype tagging SNP (htSNP) determination. We used LD Select ($r^2=0.8$) to select htSNP, selecting 146 htSNP which were typed in the remaining half of our population. No single htSNP proved associated with risk of prostate cancer. Haplotype analysis identified a 60 kb region within and flanking predicted gene *PGB* (Figure 1) associated with risk of prostate cancer in cases of families with no evidence of male-to-male transmission with two or more affected siblings (n=125) compared to our controls. Two distinct haplotypes spanning this region have been found in significant excess within these cases compared to our control population. Experiments are underway to validate this finding in a second confirmatory population of these cases.

Recently, candidate genes *SPANX-C* and *LDOC1* have been proposed for the HPC-X linkage region. We sought to confirm this finding as both of these genes lie directly upstream of *PGB*. We have assayed all database SNPs in *LDOC1* and *SPANX-C* and also identified and assayed 23 novel SNPs within and surrounding *SPANX-C*. Genotyping of *SPANX-C* proved difficult as *SPANX-C* is part of a highly homologous gene family. Therefore, using custom bioinformatics tools, we identified unique sites spanning *SPANX-C*. Using these unique sites for amplification, we then used nested amplification reactions to enable clean genotyping assays. *De novo* efforts at *LDOC1* did not result in identification of novel assayable SNPs. Using LD Select ($r^2 = 0.8$), we selected 8 htSNP for *LDOC1* and 14 htSNP for *SPANX-C*. We find no association with risk of prostate cancer for either *SPANX-C* or *LDOC1* (Tables 1 and 2).

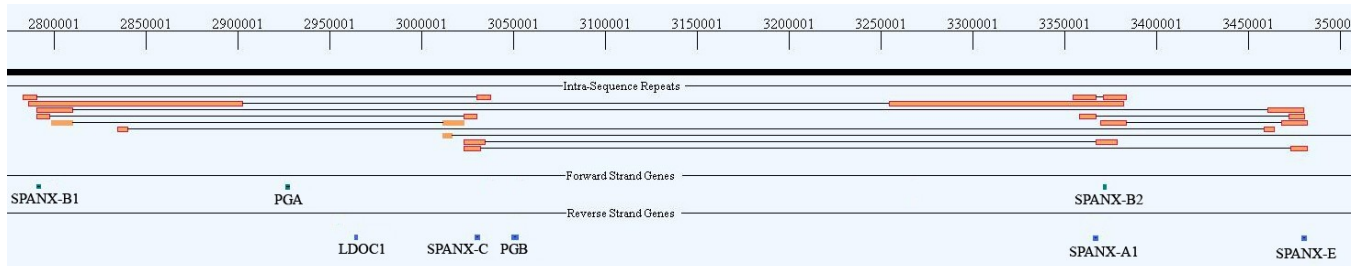


Figure 1: Graphical depiction of HPC-X region A and surrounding area covering approximately 700 kb. Note the abundance of intra-sequence repeats. Beige bars represent repeats also found elsewhere in the genome. Red-outlined beige bars represent intra-sequence repeats specific to this genomic locus. Region A includes area from 2900001-3250001 on the map above flanked by large intra-sequence repeats. Known genes present within region A are *SPANX-C* and *LDOC1*. Also shown within region A are predicted genes *PGA* and *PGB*.

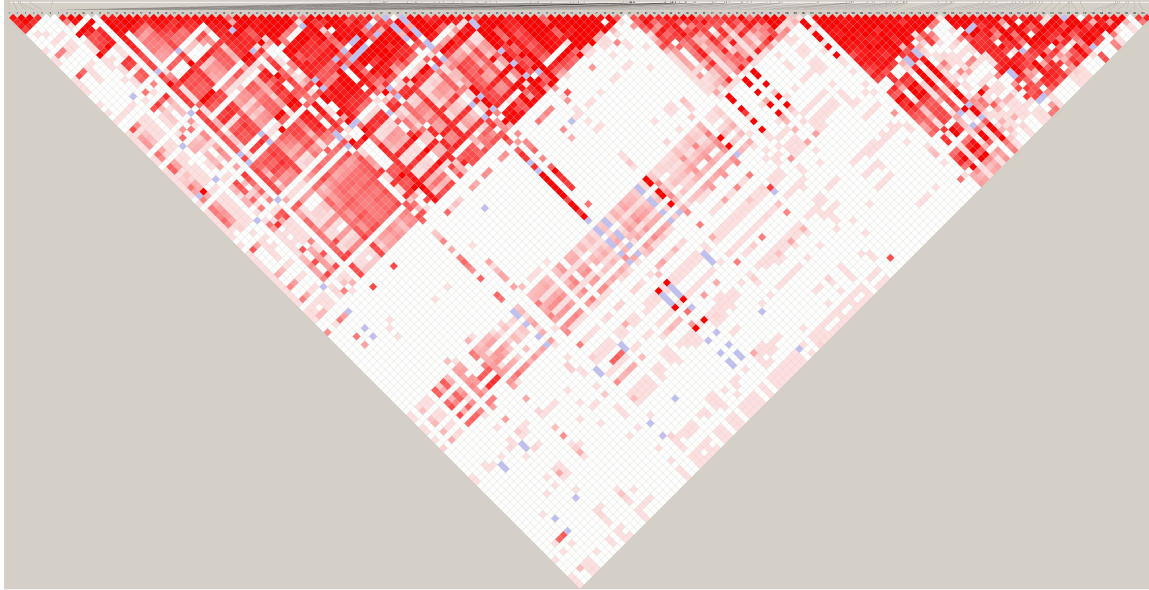


Figure 2: Linkage disequilibrium patterns at HPC-X region A for 146 htSNP in our US Caucasian population of 453 screened cancer-free controls .

Code	rs1286259	rs7883897	rs710106	rs2038375	rs4824993	rs6636233	rs3761561	rs5954228	Frequency	Odds Ratio	95% CI
1	C	G	G	G	T	C	A	C	0.47	1	reference
2	T	A	A	C	C	T	G	C	0.31	1.233	0.781 1.945
3	T	G	G	G	C	C	A	C	0.09	1.271	0.622 2.595
4	rare								0.14	0.909	0.501 1.653

Table 1: Haplotypes and frequencies for *LDOC1*. Rare haplotypes (< 2% frequency) accounted for 14% of the total genetic diversity in our US Caucasian population. This group contained an additional 20 haplotypes for a total of 23 haplotypes found in the US Caucasian population. Odds ratios and 95% CI represent results from logistic regression tests when controlling for age. Rare haplotypes were analyzed as a group. In a χ^2 test for association, the P-value for the overall haplotype profile is similarly not statistically significant (P=0.584)

Code	rs5953618	rs1059178	3029430	3029458	rs5953547	3029878	3030115	3030127	3030218	rs2144605	rs3976442	3031318	3031333	rs208264	Frequency	Odds Ratio	95% CI
1	G	C	A	G	A	A	A	C	T	T	A	G	T	C	0.34	1	reference
2	G	C	A	G	G	A	A	C	T	T	A	G	T	C	0.22	1.193	0.744 1.912
3	C	C	C	C	G	G	G	T	A	C	G	C	G	G	0.14	1.315	0.765 2.262
4	C	C	A	G	A	A	A	C	T	T	A	G	T	C	0.07	1.314	0.642 2.687
5	G	A	A	G	G	A	A	C	T	T	A	G	T	C	0.04	1.371	0.556 3.381
6	G	C	A	G	G	A	A	C	T	C	A	G	T	C	0.03	0.717	0.229 2.246
7	G	C	C	C	G	G	G	T	A	C	G	C	G	G	0.02	4.134	0.843 20.283
8	rare														0.13	1.148	0.659 2.001

Table 2: Haplotypes and frequencies for *SPANX-C*. Rare haplotypes (< 2% frequency) accounted for 13% of the total genetic diversity in our US Caucasian population. This group contained an additional 33 haplotypes, for a total of 40 haplotypes found in the US Caucasian population. Odds ratios and 95% CI represent results from logistic regression tests when controlling for age. Rare haplotypes were analyzed as a group. In a χ^2 test for association, the P-value for the overall haplotype profile is similarly not statistically significant (P=0.550)

Task 2. To determine the gene variant(s) within the associated haplotype(s) responsible for the significant association(s), and to establish a role in prostate cancer through characterization of expression in normal and transformed prostate. (Months 12-24):

- a. Sequence associated haplotype block and control to identify possible missed polymorphisms in the associated haplotype
- b. Type new high-impact polymorphisms in the population to determine frequency in the study population and rule out possibility of private mutation
- c. Determine expression patterns of candidate gene in normal and transformed prostate tissue, and in prostate samples from patients with the disease-associated variant and without using Northern blot analysis and RT-PCR

Task 2a is currently in progress. As previously mentioned we have identified a 60 kb region within and flanking *PGB* associated with risk of prostate cancer. We are currently re-sequencing predicted exons as well as conserved non-coding areas using representative samples of our population for each haplotype. Members will be selected in duplicate for each haplotype over the 60 kb region. Tasks 2b and 2c are contingent on completion of task 2a and will be worked on following completion of task 2a.

Task 3. To confirm associated prostate cancer gene variants in a second study population currently under ascertainment, and to extend investigation in an African American Study population. (Months 24-36)

- a. Identify previously unknown polymorphism via SSCP in our confirmatory population in genomic locus HPC-X
- b. Genotype population for polymorphism
- c. Analyze genotypes to determine genetic architecture and haplotypic diversity of HPC-X for comparison with the Caucasian study population

Task 3a-c is in preparation and will be completed upon analysis of data obtained in task 2. In the meantime, we are actively increasing our ascertainment efforts. To date we have 146 African American cases and 93 African American controls.

Key Research Accomplishments

1. Ascertainment of a US Caucasian population with statistical power to detect common variants that may predispose prostate cancer risk.
2. *De novo* SNP discovery leading to discovery of 66 unpublished SNPs in a US Caucasian population

3. No association with risk of prostate cancer for genes *LDOC1* and *SPANX-C* in our study population.
4. Preliminary identification of a 60 kb area within and flanking predicted gene *PGB* associated with risk of prostate cancer. This finding is currently undergoing replication in our lab.

Reportable Outcomes

No association with risk of prostate cancer for *LDOC1* and *SPANX-C* candidate genes within the HPC-X locus in a US Caucasian study population

Poster presentation at the American Society of Human Genetics Meeting, New Orleans, LA 2006

Yaspan B, Elmore JB, Breyer J, Bradley K, McReynolds K, Smith JR

Conclusions

In the past year, we have begun to systematically dissect HPC-X to uncover the variant or variants responsible for its association with risk of prostate cancer. Starting with region A, whose boundaries we identified through shared haplotype analysis of founder populations, we have identified 2 haplotypes spanning a 60 kb region associated with risk of prostate cancer in and around a predicted gene, *PGB*. Although located directly upstream of *PGB*, we found no evidence that either *SPANX-C* or *LDOC1* are associated with risk of prostate cancer.

Future Directions

In the next year, in addition to accomplishing Task 2 for region A, we will begin dissecting region B, the boundaries of which were defined by preliminary data of shared haplotype analysis in founder populations. Since the inception of this project, the International HapMap Consortium has released version 2.0 of the HapMap. We intend to utilize this resource for selecting htSNPs across region B for typing in our population by looking at the htSNP ($r^2 = 0.80$) selected for the CEPH population (Utah residents with ancestry from northern and western Europe). We will select htSNP for genes identified in public databases and potential coding regions identified using custom bioinformatics software. We will then genotype this set of htSNP in our study population.